

Figure 1

The novel gene as identified through RACE analysis (894 bp)

GGGAGTGGAGTGAGGGGTAACAAGATGGCGACCGAGACGGTGGAGCTCCATAAGCTAA
AGCTTGCCGAACATAAGCAAGAATGCTTGCTCGTGGTTTGGAGACCAAGGGAATAAAG
CAAGATCTTATCCACAGACTCCAGGCATATCTTGAAGAACATGCTGAAGAGGAGGCAAAT
GAAGAAGATGTACTGGGAGATGAAACAGAGGAAGAAGAAACAAAGCCCATTGAGCTCCC
TGTCAAAGAGGAAGAACCCCTGAAAAAAGTGTGATGTGGCAGCAGAGAAGAAAGTGG
TGAAAATTACATCTGAAATACCACAGACTGAGAGAATGCAGAAGAGGGCTGAACGATTCA
ATGTACCTGTGAGCTTGGAGAGTAAGAAAGCTGCTCGGGCAGCTAGGTTTGGGATTCT
TCAGTTCCAACAAAAGGCTCTGTCATCTGATAACAAACCTATGGTTAACTTGGATAAGCTG
AAGGAAAGAGCTCAAAGATTTGGTTTGAATGTCTCTTCAATCTCCAGAAAGTCTGAAGAT
GATGAGAAACTGAAAAAGAGGAAGGAGCGATTTGGGATTGTCAAGTTCAGCTGGAAC
TGGAACACAGAGGATACAGAGGCAAGAGAGGAAAAGAGCAGAGCGCTTTGGGATT
GCCTGATGAAAAGTTCCTGATACTTTCTGTTCTCCAGTGTTCCTCATTTCTCCTTCTTC
TTGGTCACATATATGCCTAAATGCACAGTCATGTGCCTACGTCCTCGCTCGCAATGAGG
GAGCATGTACCCAGGTACATCCATGAAGTGCAGCAGAGTTGACTTATTGCTGTTTCA
GCTTTAAGGTTGTTGTGTTTTGTTTTGATTATGTTGCTTGTTAATAAAAAAAATAGAAA

A

03783676-02240

Figure 2

Amino acid sequence as translated from the novel gene (210 amino acids)

MATETVELHKLKLAELKQECLARGLETGKIQDLIHRLQAYLEEHAEEDVVGDETEEE
ETKPIELPVKEEPEKTVDAEKKVVKITSEIPQTERMQKRAERFNPVVSLESKKAARAAR
FGISSVPTKGLSSDNKPMVNLDKLKERARFGLNVSSISRKSEDEKLKKRKERFGIVTSSAG
TGTTEDEAKKRKRAERFGIA

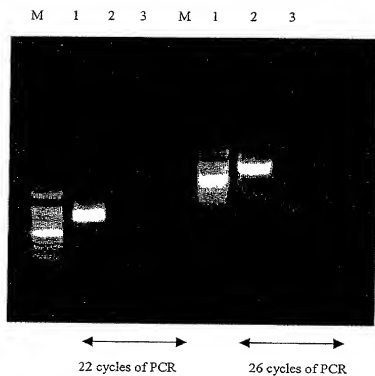
Underlined sequences are amino acid sequences obtained by MS/MS analysis.

Figure 3

The sequence of the novel gene amplified through long distant PCR and used to construct the expression vector (873 bp).

TGGAGTGAGGGGTAACAAGATGGCGACCGAGACGGTGGAGCTCCATAAGCTAAAGCTT
GCCGAACATAAGCAAGAATGTCTTGCTCGTGTTGGAGACCAAGGGAATAAAGCAAGA
TCTTATCCACAGACTCCAGGCATATCTTGAAGAACATGCTGAAGAGGAGGCAAATGAAG
AAGATGTACTGGGAGATGAAACAGAGGAAGAAGAAACAAAGCCCATTGAGCTCCCTGTC
AAAGAGGAAGAACCCCTGAAAAAACTGTTGATGTGGCAGCAGAGAAGAAAGTGGTGAA
AATTACATCTGAAATACCACAGACTGAGAGAATGCAGAAGAGGGCTGAACGATTCAATGT
ACCTGTGAGCTTGGAGAGTAAGAAAGCTGCTCGGGCAGCTAGGTTTGGGATTTCTTCAG
TTCCAACAAAAGGTCTGTCATCTGATAACAAACCTATGGTTAACTTGGATAAGCTGAAGG
AAAGAGCTCAAAGATTTGGTTTGAATGTCTCTTCAATCTCCAGAAAGTCTGAAGATGATG
AGAAACTGAAAAGAGGAAGGAGCGATTTGGGATTGTCAAGTTTCAGCTGGAAC TGGA
ACCACAGAGGATACAGAGGCAAGAGAGGAAAAGAGCAGAGCGCTTTGGGATTGCCT
GATGAAAAGTTCCTGATACTTTCTGTTCTCCAGTGTCTTCCATTTCTCTCCTTCTTCTTGG
TCACATATATGCCTAAATGCACAGTCATGTGCCTACGTCCTGCCTCGCAATGAGGGAGC
ATGTACCCACAGGTACATCCATGAAGTGCAGGAGCAGTTTGACTTATTGCTGTTTCAGCTT
TAAGGTTGTTGTGTTTTTGTGTTTGAATTATGTTGCTTGTTAAAT

Figure 4



00760426-002101

Figure 5

P-151 5'-Untranslated Region

```
1                                     75
CAGGGGCAGCAGTGAATTATCTGAACTCGGATCTTTAAAATGTGGTAGCTCTAAAGCTGATGATGTCTGGTTAGG
*****
76                                     150
AAGTGGCTCTTGCCCGCCCCAGCCCCACCGCAGTTCCTTAAGCCCCGCCCATGCCCTCCAGCTTCCTCCTCA
*****
151                                     225
TGTTCATCGGTTTTTTCAGGGCTCCCTTCAACGCTCCCTCTCAGTATTTAGGTCACCACTCCCTCGGCGCCCTT
*****
226                                     300
TTCGCCTCCCACTTTTTCCTCAGCAACCTTACAGTCTTTGCAGCTCCTACCTGCCAGCTCAGATCCCGCTCC
*****
301                                     375
GGCTATGGGCGCGGCGCGGCTACCAACCTGAAGTCTCCAGGAAGTAAAGCCTCTCCTTCTGCCCTTTCCTGT
376                                     450
TGGAGGAACAGAATCAGCGCTGCCACCACCCATTGGTTGGTGGTCTGTAAATGCAGAAGCACAGTTGGTTGCCATT
451                                     525
TCTGTCGTTTCGAAGATACAGTGCCCGCCCTCTCCAGTTCCACCTTTTGAAAGAGGTGGGGCAAGCTGCCTAG
526                                     600
AGAAGTGAAGAGCGACGTGAGCTATTGACCAATGGGAAGAGCTGATGGTATGGCGTGGGAGCAAGAGTGACAACGA
601                                     675
TTGGTCAGCCTTGACATCTCTACGCCTAAGCGGGAACTCCTGGAGGCGGAGCCCGGGTGGGGGGAGTGGAGTG
676
AGGGGTAAACAAGATG.....P151 coding region.....
```

(Total length: 690 bp)

Sequence with asterisk: the 274 bp fragment

Underlined sequences are the minicistrons or uORFs before the start of the P151 coding region with the start and stop codons in bold.